

Appendix A

Quality Criteria 1	
QC1.1	Match the Standards
	Items/Tasks match the academic content standards-state-reading
	Items/Tasks match the academic content standards-state-mathematics
	Items/Tasks match the academic content standards-state-science
QC1.2	Pattern of Emphasis
	Items/Tasks reflect the content standards and process pattern-state-reading
	Items/Tasks reflect the content standards and process pattern-state-mathematics
	Items/Tasks reflect the content standards and process pattern-state-science
QC1.3	Depth of Knowledge
	Items/Tasks are distributed across each performance level-reading
	Items/Tasks are distributed across each performance level-mathematics
	Items/Tasks are distributed across each performance level-science
	Response formats are varied
	Items/Tasks are focused on higher order thinking skills
QC1.4	Range of Knowledge
	Items/Tasks are distributed across the academic content standards-reading
	Items/Tasks are distributed across the academic content standards-mathematics
	Items/Tasks are distributed across the academic content standards-science
	Scores reflect the full range of achievement-reading
	Scores reflect the full range of achievement-mathematics
	Scores reflect the full range of achievement-science
QA 1	Quality Assurance
	Alignment reviews are conducted by internal groups
	Alignment reviews are conducted by external groups
	Alignment results are documented
	Alignment results are used in future item development
	Alignment training is used to develop future assessment literacy

Quality Criteria 2	
QC2.1	Administering
	Setting, time, and accommodations parameters of the assessments are articulated
	Instructions to students articulate behavioral expectations
	Test security of procedures ensure assessments are not compromised
	Inclusion guidelines articulate how students should participate
	Non-standard administration qualifiers and score reporting are concise
	Timelines, logistics, and other procedural expectations are established
QC2.2	Student-level Reporting
	Individual student reports are sent to parents prior to the upcoming school year
	Student achievement reported by performance levels for reading and mathematics
	Student achievement reported using the adopted content-based competency descriptors
	Performance scores are reported to reflect SEM ranges
	Reports provide explanatory narratives for parents and students
	Reports are provided in additional language when appropriate
	Reports are disseminated to ensure student confidentiality
	Reports provide subdomain results when appropriate
QC2.3	School/District level Reporting
	School/District level participation results aggregated by gender, ethnicity, ED, SWD, LEP, Migrant, Alternate
	School/District level assessment results aggregated by gender, ethnicity, ED, SWD, LEP, Migrant, Alternate
	School-level reports reflect achievement across the PLs
	Reports provide explanatory narratives for parents and students
	Reports are disseminated to stakeholders in a timely manner
	Reports prevent the identification of individual student results
QC2.4	Participation Rates
	Enrollment-Participation rates account for all students-aggregate
	Enrollment-Participation rates account for all students-subgroups
	Policies require all students to participate in the assessment system
	Guidelines for determining how SWD should participate in the assessment system
	Guidelines for determining how ELL should participate in the assessment system
QC 2	Quality Assurance
	Administrative guidelines are documented for each component of the assessment system
	Monitoring/verification procedures are established to ensure inclusion, standardization, and security
	Training is provided to individuals administering the assessments
	Reported data are verified prior to public release
	Quality criteria ensure reporting elements are documented
	Training is provided for score interpretation and instructional use
	Internal auditing procedures ensure confidential student information is not released
	Public reporting timelines and formats are clearly defined.
	Accommodations used during assessment are cross-referenced with a student's IEP
	Alternate assessment participation trend data is examined annually

Quality Criteria 3	
QC3.1	Items/Tasks Development
	Items/Tasks are designed within specified blueprints or specification tables
	Items/Tasks specifications consider difficulty and discrimination properties
	Items/Tasks address higher order thinking skills using a variety of response options
	Items/Tasks are developmentally appropriate
	Items/Tasks are focused on higher order thinking skills
QC3.2	Operational Forms
	Operational forms have been reviewed to minimize structural confounding influences
	Operational forms have been reviewed for bias and sensitivity
	Operational forms have been evaluated against Universal Design principles
	Operational forms have been comprehensively edited
	Operational forms administration allows for appropriate accommodations
	Operational forms contain items/tasks being field tested for future use
QC3	Quality Assurance
	Operational forms are updated following an on-going plan to improve test quality
	Reviewers, criteria, and recommendations are documented
	Operational forms being used within the current assessment system are documented
	Operational form modifications or comprehensive review follow a prescribed schedule
	Design and appropriateness trainings are used to build capacity

Quality Criteria 4	
QC4.1	Performance Level Descriptors
	PLDs describe the achievement continuum using content-based competencies-reading
	PLDs describe the achievement continuum using content-based competencies-mathematics
	PLDs describe the achievement continuum using content-based competencies-science
	Cut scores are established for each performance level-reading
	Cut scores are established for each performance level-mathematics
	Cut scores are established for each performance level-science
QC4.2	Performance Level Calibration
	Performance levels (cut scores) within each grade level are comparable to ensure augmentation of scores -reading
	Performance levels (cut scores) across grade level are comparable to ensure augmentation of scores - reading
	Performance levels (cut scores) within each grade level are comparable to ensure augmentation of scores -mathematics
	Performance levels (cut scores) across grade level are comparable to ensure augmentation of scores - mathematics
	Performance levels (cut scores) within each grade level are comparable to ensure augmentation of scores -science
	Performance levels (cut scores) across grade level are comparable to ensure augmentation of scores - science

QC4.3	Scoring
	Machine scored responses have item/task scoring guidelines
	Machine scored response documents have guidelines for reporting demographic information
	Machine scored responses have clearly defined business rules for data integrity, aggregation, and PL assignment
	Hand-scored responses have item/task scoring guidelines
	Hand-scored documents have guidelines for reporting demographic information
	Hand-scored responses have clearly defined business rules for data integrity, aggregation, and PL assignment
QC4.4	Reliability Evidence
	Methods for determining score consistency are technically appropriate
	Reliability coefficients are reported for all assessments
	Standard and conditional errors of measure are reported
	Decision consistency across the performance levels are within acceptable ranges
	Generalizability evidence associated with internal consistency of item/task responses or others is reported
	Validity Evidence
	Purpose statements are clear and focus on student achievement
	Intended consequences are evaluated each year
	Evidence demonstrates items/tasks are measuring the knowledge and skills within the standards
	Inappropriate score inferences are minimized
	End-user feedback
	Evidence documents the intended cognitive processes are being evaluated
	Scoring structures and item interrelationships are consistent with the assessment design
	Evidence documents the test characteristics are within expected parameters
	Evidence documents how confounding variables are eliminated or minimized
	Descriptive statistics examining assessment results
	Item-level difficulty and discriminatory properties
	Item calibration and scale development
QC4	Quality Assurance
	PLDs were developed and reviewed by a representative group of stakeholders
	PLD "cut scores" were established using technically recognized procedures
	PLDs with associated cut score ranges were disseminated to educators and parents
	Standard setting procedures are documented in sufficient detail for external verification
	PL calibrations followed technically recognized procedures
	Machine scoring procedures are comprehensively documented
	Hand-scoring procedures are comprehensively documented
	Data quality procedures are implemented prior to score aggregation and reporting
	Reliability results are within acceptable ranges
	Assessment quality improvements are linked to validity/reliability evidence
	Technical criteria are documented after each administration
	Validity evidence procedures are technically appropriate
	Assessment quality improvements are linked to validity/reliability evidence